

**Controls protocol.** Control subjects have been and continue to be recruited by Knowledge Networks (San Jose, CA), a survey and marketing research company. The KN web-enabled panel is a method for conducting Internet-based survey research with a U.S. representative probability sample. The KN Panel, recruited randomly through Random Digit Dialing, represents key demographic dimensions of the U.S. population. The web-enabled panel tracks closely the U.S. population on age, race, Hispanic ethnicity, geographical region, employment status, and other demographic elements. The differences that do exist are corrected statistically in survey data (i.e., by non-response adjustments). The web-enabled panel is comprised of both Internet and non-Internet households, all of which are provided the same equipment for participation in Internet surveys. There are four main factors responsible for the representativeness of the web-enabled research panel. First, the panel sample is selected using list-assisted random digit dialing telephone methodology, providing a probability-based starting sample of U.S. telephone households. Second, the panel sample weights are adjusted to U.S. Census demographic benchmarks to reduce error due to noncoverage of nontelephone households and to reduce bias due to nonresponse and other nonsampling errors. Third, samples selected from the panel for individual studies are selected using probability methods. Appropriate sample design weights for each study are calculated based on specific design parameters (the probability of each panelist's being selected into the sample for the individual study so that the selected sample will be similar to the benchmarks for the relevant populations). Fourth, nonresponse and poststratification weighting adjustments are applied to the final survey data to reduce the effects of nonsampling error (variance and bias -- once the data collection is complete).

KN provides households in the panel with free Web access and an Internet appliance. In return, panel members participate in 10- to 15-minute Internet surveys three to four times a month. The panel does not respond significantly differently over time to surveys than more "naïve" survey respondents with less tenure on the panel (Dennis, 2001). Survey responses are confidential, with identifying information never revealed without respondent approval. When surveys are assigned to panel members, they receive notice in their password protected e-mail account that the survey is available for completion. Surveys are self-administered and accessible any time of day for a designated period. Participants can complete a survey only once. Members may leave the panel at any time, and receipt of the WebTV and Internet service is not contingent on completion of any particular survey.

The follow table compares estimates from the Knowledge Networks Panel to statistics from the 2003 Current Population Survey.

**Table 1: Demographics of Knowledge Networks Panel - January 2004**

		Knowledge Networks Adult Panel	Adult U.S. Population (November 2003 CPS)
<b>Gender</b>	Male	48.1%	48.0%
	Female	51.9%	52.0%
<b>Age</b>	18-24	11.3%	12.8%
	25-34	18.5%	18.3%
	35-44	22.1%	20.4%
	45-54	19.4%	19.1%
	55-64	14.5%	13.2%
	65 or over	14.3%	16.1%
<b>Race</b>	White Only	79.7%	82.1%
	Black Only	12.4%	11.5%
	American Indian, Alaskan Native Only	1.5%	0.7%
	Asian Only	3.2%	4.2%
	Hawaiian or Pacific Islander Only	0.2%	0.2%
	2+ Races	3.1%	1.3%
<b>Hispanic Ethnicity</b>	Hispanic	10.9%	12.5%
	Non-Hispanic	89.1%	87.5%
<b>Employment Status</b>	In the Labor Force	62.7%	65.8%
	<i>Working full-time</i>	49.9%	55.3%
	<i>Working part-time</i>	12.8%	10.6%
	Not in the Labor Force	37.3%	34.2%
<b>Marital Status</b>	Married	58.9%	57.7%
	Not married	41.1%	42.3%
<b>Housing Ownership</b>	Own	65.9%	67.5%
	Rent/Other	34.1%	32.5%
<b>Level of Education</b>	Less than High School Diploma	15.2%	15.5%
	High School Diploma or Equivalent	32.3%	32.4%
	Some College	21.4%	19.1%
	Associate Degree	6.6%	7.7%
	Bachelor's Degree or Beyond	24.5%	25.3%
<b>Household Income</b>	Under \$10,000	8.3%	7.5%
	\$10,000-\$24,999	18.8%	17.9%
	\$25,000-\$49,999	34.0%	28.6%
	\$50,000-\$74,999	20.5%	19.9%
	\$75,000 or more	18.4%	26.2%
<b>Census Region</b>	Northeast	18.6%	19.3%
	Midwest	22.2%	22.7%
	South	36.1%	35.4%
	West	23.0%	22.5%
<b>Internet Access</b>	Any Connection Speed	59.7%	56.0%
	Broadband	17.2%	18.0%

KN has completed the EA sample of 3,297 controls and it has partnered with Sampling International ([www.surveypot.com](http://www.surveypot.com)), a company with a large pool of African American panelists, to help complete the recruitment of 1,200 AA controls, which is expected to be finalized in August of 2006. All control subjects complete an online, short self-report clinical assessment after giving informed consent through an online procedure, and prior to venipuncture being arranged. The consent includes an authorization for qualified scientists, with approval from NIMH, to study any condition or trait. Participants are asked their age, sex, height, weight, the ethnic background of their grandparents (which is helpful for genetic studies), and questions about certain attitudes, behaviors, feelings, life experiences, and mental health or substance use problems. People with a history of mental health or substance abuse problems are eligible to participate in this study in the control group. Therefore, the rationale for the assessment procedure was as follows: these control DNAs will be used to study a wide range of psychiatric disorders (in addition to non-psychiatric disorders). Investigators accessing this collection might wish to select a control group from individuals who are less likely to have suffered from the disorder under study. For the most common psychiatric disorders, such as depression, anxiety, and substance use disorders, screening out some substantial proportion of affected individuals could conceivably increase the power to detect association. Also, self-report of these more common disorders is reasonably sensitive and specific. Therefore, we adopted a self-report screening procedure focused on common psychiatric disorders, which can be completed in about 20 minutes. We also added brief assessments of several variables of interest to the field (see below), and a very brief inventory of personality variables that have been of wide interest (neuroticism and extraversion).

The controls are assessed with:

- a. CIDI-SF-lifetime/online. This is an online, modified version of the Composite Instrument for Diagnostic Interviewing (CIDI)-Short Form v1.1 Dec 2002 (World Health Organization, [http://www3.who.int/cidi/cidi-sf\\_\\_12-03-02.pdf](http://www3.who.int/cidi/cidi-sf__12-03-02.pdf)), including sections for major depression, generalized anxiety, specific phobia, panic attack, agoraphobia, social phobia, alcohol dependence and drug dependence. This version of the CIDI was originally worded for symptoms and disorder during the past 12 months. After consulting with participants in the development of the CIDI, we modified the wording of each section to query lifetime episodes -- note that the CIDI was originally validated using lifetime diagnoses. There is no "official" WHO endorsement of modified versions of the CIDI.
- b. A brief inventory for nicotine dependence (Caryn Lerman, Ph.D., personal communication).
- c. A question about sexual orientation (Michael Bayley, personal communication).
- d. Single items about lifetime diagnosis of or treatment for psychotic or bipolar disorders.
- e. The 12-item versions of the Eysenck neuroticism and extraversion scales (Eysenck et al., 1985).
- f. Current height, and highest lifetime weight.

There are 69 screening questions (45 for disease related traits and 24 for personality traits). In preliminary analyses of 3,297 EA controls, we have detected and excluded several small sets of outliers: 12 subjects that have answered "yes" to 50 or more screening questions, 12 subjects who refused to answer 4 or more of the screening questions, and 19 subjects excluded due to software failures. Around 7.85% of controls were subsequently excluded from association experiments because they have endorsed lifetime history of a psychotic or bipolar disorder or of any hallucinations or delusions or who have failed to answer these questions.

Only KN and the phlebotomy contractor (EMSI) see the link between subject name and ID, and after one year (to permit subject withdrawal), the link is destroyed to anonymize the sample. All controls have consent in writing to use of their DNA for study of any medical disorder.